

Roteiro Trabalho Final – Modelos de Crescimento Econômico

Ao longo do curso e especialmente na 1ª aula, vimos uma série de fontes de dados econômicos que podem ser usadas nos modelos estudados, especialmente para o Brasil. O objetivo deste trabalho é construir um *dataset* com dados de inflação e de emprego para estimar uma Curva de Phillips para o Brasil.

1. Instale a biblioteca *ipeadatapy* e atualize o *statsmodels* (ver aula da Curva de Phillips para ver como fazer isso).
2. Usando o *ipeadatapy*, colete a taxa de inflação mensal anualizada para o Brasil com o comando `dfIPCA = ipeadatapy.timeseries('PAN12_IPCAG12')`
 - a. Note que os valores da série estão multiplicados por 100 e em valores anualizados. Primeiro, divida a série por 100 (para ter os valores no formato numérico, e não percentual) e mensalize a série.
 - i. Dica: fizemos isso quando lemos os dados do Tesouro Americano na aula de CAPM!
 - b. Depois de mensalizar a série, multiplique por 100 para ter os valores novamente em percentuais.
 - c. Para deixar os dados limpos, renomeie a coluna de "VALUE ((% a.a.))" para "IPCA" e deixe apenas ela no DataFrame (`dfIPCA = dfIPCA[["VALUE ((% a.a.))"].copy()`).
3. Usando o *ipeadatapy*, colete a série de desemprego para o Brasil. Aqui, haverá um desafio: será preciso unir duas séries de desemprego, o que será feito mais a frente.
 - a. Usando o comando `dfDesemprego = ipeadatapy.timeseries('PAN12_TD12')`, colete a série da taxa de desemprego do IBGE que vai de março de 2002 a dezembro de 2015. Para deixar os dados limpos, renomeie a coluna de "VALUE ((%))" para "Desemprego" e deixe apenas ela no DataFrame (`dfDesemprego = dfDesemprego[["VALUE ((%)")].copy()`).
 - b. Rode `dfDesocupacao = ipeadatapy.timeseries('PAN12_TDESOC12')` e colete a série da taxa de desocupação do IBGE que vai de março de 2012 ao presente. Para deixar os dados limpos, renomeie a coluna de "VALUE ((%))" para "Desocupacao" e deixe apenas ela no DataFrame (`dfDesocupacao = dfDesocupacao[["VALUE ((%)")].copy()`).
4. Crie um dataset que contenha três colunas, uma para cada um dos dados coletados acima (além do *index* contendo as datas). Para isso, use o método *merge* do Pandas, mantendo como opção *how="left"* para performar um *left join*. Chame o DataFrame resultante de *dfDados*.
 - i. Dica: para manter o máximo de informação possível, comece com os dados do IPCA (`dfIPCA.merge(dfDesemprego.....)`)

5. Usando o `np.where()`, crie uma coluna unificada para a taxa de desemprego, chamando-a de "Taxa_Desemprego". A regra para a unificação é: se houver dados de "Desocupacao", use-os; caso contrário, mantenha os valores de "Desemprego".
6. Selecione apenas as colunas de "IPCA" e "Taxa_Desemprego" no `dfDados` e remova as linhas com valores nulos, mantendo apenas os meses com informações para ambas as variáveis.
7. Usando o `regplot` do `seaborn`, plote a relação entre "IPCA" e "Taxa_Desemprego". A Curva de Phillips tradicional parece estar valendo para o Brasil no período analisado? Lembre-se, por esse modelo, espera-se que haja uma relação negativa entre essas variáveis.
8. Execute o modelo de Phillips levando em conta expectativas adaptativas.
 - a. Usando o método `.diff()`, tire as diferenças da taxa de inflação ("IPCA") e da taxa de desemprego ("Taxa_Desemprego"). Depois, drope a 1ª linha (terá valores nulos). Note que as variáveis criadas são justamente a variação da taxa de inflação e a variação da taxa de desemprego.
 - b. Teste a estacionariedade da variação da inflação e da variação do desemprego usando a função `adfuller` do `statsmodels`. Ambas as séries são estacionárias?
 - c. Verifique se a variação da inflação possui autocorrelação parcial consigo mesma, usando o `plot_pacf()`, ou seja, se o modelo de Expectativas Adaptativas realmente é válido.
 - i. Dica: para ser válido, o modelo precisa que ao menos uma autocorrelação parcial seja estatisticamente diferente de 0.
 - d. Usando a função `ARIMA()`, rode um modelo AR(1) para a variação da taxa de inflação, tendo como variável exógena a variação da taxa de desemprego. A relação da Curva de Phillips (coeficiente da variação da taxa de desemprego) é significativa a **15%**? Ademais, o modelo sugere que há persistência inflacionária (componente autorregressivo > 0)?
 - e. Analisando o sumário e os diagnósticos do modelo (`plot_diagnostics`), analise se o resíduo do AR(1) é um ruído branco. Para isso, ele deve ser homoscedástico, ter distribuição normal e não possuir autocorrelação. Com base nisso, responda: o modelo pode ser melhorado? Lembre-se que o objetivo de qualquer modelo temporal é conseguir produzir um resíduo que seja um ruído branco.